

FIG. 1

gattctcagt agagacgttt gactgtccca acccgatgct gccttccac ataaatgaga 60

tttttttctg ccaggcaac atg gtt tta ccc tca tat tca aaa aaa ccc tta 112
Met Val Leu Pro Ser Tyr Ser Lys Lys Pro Leu
1 5 10

atc tct aat gtg gag cag ctg atc ctg ggg atc ccg ggc cag aat cgc 160
Ile Ser Asn Val Glu Gln Leu Ile Leu Gly Ile Pro Gly Gln Asn Arg
15 20 25

cgg gag ata ggc cat ggc cag gat atc ttt cca gca gag aag ctc tgc 208
Arg Glu Ile Gly His Gly Gln Asp Ile Phe Pro Ala Glu Lys Leu Cys
30 35 40

cat ctg cag gat cgc aag gtg aac ctt cac aga gct gcc tgg ggc gag 256
His Leu Gln Asp Arg Lys Val Asn Leu His Arg Ala Ala Trp Gly Glu
45 50 55

tgt att gtt gca ccc aag act ctc agc ttc tct tac tgt cag ggg acc 304
Cys Ile Val Ala Pro Lys Thr Leu Ser Phe Ser Tyr Cys Gln Gly Thr
60 65 70 75

tgc ccg gcc ctc aac agt gag ctc cgt cat tcc agc ttt gag tgc tat 352
Cys Pro Ala Leu Asn Ser Glu Leu Arg His Ser Ser Phe Glu Cys Tyr
80 85 90

aag agg gca gta cct acc tgt ccc tgg ctc ttc cag acc tgc cgt ccc 400
Lys Arg Ala Val Pro Thr Cys Pro Trp Leu Phe Gln Thr Cys Arg Pro
95 100 105

acc atg gtc aga ctc ttc tcc ctg atg gtc cag gat gac gaa cac aag 448
Thr Met Val Arg Leu Phe Ser Leu Met Val Gln Asp Asp Glu His Lys
110 115 120

atg agt gtg cac tat gtg aac act tcc ttg gtg gag aag tgt ggc tgc 496
Met Ser Val His Tyr Val Asn Thr Ser Leu Val Glu Lys Cys Gly Cys
125 130 135

tct tga gataccccaa agcctcctac tggcctcagg gccacctaag tctcaggact 552
Ser
140

ttagtagggg gtgggattac ttttcatagc aagtagagct ctttgaaggg aggtgggatt 612

tggtttgttt ctcaaagcac agcaagaagg ttggcattat ggcagtaaca aat 665

FIG. 2A

actagtgatt ctccagtagag acgttttgact gtcccaaccc gatgctgcct tcccacataa 60

atg aga ttt ttt tct gcc agg caa cat ggt ttt acc ctc ata ttc aaa 108
Met Arg Phe Phe Ser Ala Arg Gln His Gly Phe Thr Leu Ile Phe Lys
1 5 10 15

aag aca aag att cca gcc act gat gtc gct gat gcc agc ctg aat gaa 156
Lys Thr Lys Ile Pro Ala Thr Asp Val Ala Asp Ala Ser Leu Asn Glu
20 25 30

tgt tcc agt acc gaa agg aaa caa gac gta gtg ttg ctg ttc gtg acc 204
Cys Ser Ser Thr Glu Arg Lys Gln Asp Val Val Leu Leu Phe Val Thr
35 40 45

ttg tcc cac aca cag cca cct ctg ttt cac ctg cct tat gtc cag aaa 252
Leu Ser His Thr Gln Pro Pro Leu Phe His Leu Pro Tyr Val Gln Lys
50 55 60

ccc tta atc tct aat gtg gag cag ctg atc ctg ggg atc ccg ggc cag 300
Pro Leu Ile Ser Asn Val Glu Gln Leu Ile Leu Gly Ile Pro Gly Gln
65 70 75 80

aat cgc cgg gag ata ggc cat ggc cag gat atc ttt cca gca gag aag 348
Asn Arg Arg Glu Ile Gly His Gly Gln Asp Ile Phe Pro Ala Glu Lys
85 90 95

ctc tgc cat ctg cag gat cgc aag gtg aac ctt cac aga gct gcc tgg 396
Leu Cys His Leu Gln Asp Arg Lys Val Asn Leu His Arg Ala Ala Trp
100 105 110

ggc gag tgt att gtt gca ccc aag act ctc agc ttc tct tac tgt cag 444
Gly Glu Cys Ile Val Ala Pro Lys Thr Leu Ser Phe Ser Tyr Cys Gln
115 120 125

ggg acc tgc ccg gcc ctc aac agt gag ctc cgt cat tcc agc ttt gag 492
Gly Thr Cys Pro Ala Leu Asn Ser Glu Leu Arg His Ser Ser Phe Glu
130 135 140

tgc tat aag agg gca gta cct acc tgt ccc tgg ctc ttc cag acc tgc 540
Cys Tyr Lys Arg Ala Val Pro Thr Cys Pro Trp Leu Phe Gln Thr Cys
145 150 155 160

cgt ccc acc atg gtc aga ctc ttc tcc ctg atg gtc cag gat gac gaa 588
Arg Pro Thr Met Val Arg Leu Phe Ser Leu Met Val Gln Asp Asp Glu
165 170 175

cac aag atg agt gtg cac tat gtg aac act tcc ttg gtg gag aag tgt 636
His Lys Met Ser Val His Tyr Val Asn Thr Ser Leu Val Glu Lys Cys
180 185 190

ggc tgc tct tga gataccccaa agcctcctac tggcctcagg gccacctaag 688
Gly Cys Ser
195

U.S. Pat. & Tm. Off.

FIG. 2B

tctcaggact ttagtagggg gtgggattac ttttcatagc aagtagagct ctttgaaggg 748
aggtgggatt tggtttggtt ctcaaagcac agcaagaagg ttggcattat ggcagtaaaa 808
tc 810

FIG. 2B

FIG. 3

```

201 FLEILVKEDRD SGVNFQPEDTCARLRCSLHASLLVVTLNPDQC...HPSR 247
      :      :.   ... | :|      | .
1 .....MVLPSYSKKPLIS.NVEQLILGIPGQ 25

248 KRRAAIPVPKL.SCKNLCHRHQLFINFRDLGWHKWI IAPKGFMANYCHGE 296
    ||      :   . |||      :|      | . |:|||      .|||
26 NRREIGHGQDIFPAEKLCHLQDRKVN LHRAAWGECIVAPKTL SFSYCQGT 75

297 CPFSLTISLNSSNYAFMQALMHA VDPEIPQ..AVCIPTKLSPISMLYQDN 344
    || .|      | |.:      :   || |      | || .      |.: ||.
76 CP.ALNSEL RHSSF...ECYKRAV.PTCPWLFQTCRPTMVRLFSLMVQDD 120

345 NDNVILRHYEDMVVDECGCG 364
      . . :      .|: .|||
121 EHKMSVHYVNTSLVEKCGCS 140

```

Percent Similarity: 36.567 Percent Identity: 26.866

FIG. 4

```

151 QEPHVWGQTTPKPGKMFVLRSPWPQGAVHFNLLDVAKDWNNDNPRKNFGL 200
      : |
1  .....MRFFSARQHGF 11
201 FLEILVKEDRDSGVNFQPEDTCARLRCSLHASLLVVTLN...PDQCH... 244
      | . . | . | . || || . | |
12 TLIFKKTKIPATDVADASLNECSSTERKQDVLLFVTLSTQPPPLFHLPY 61
245 ...P..SRKRRAAIPVPK.....LSCKNLCHRHQLFINFRDLG 277
      | | . : : | . || | : |
62 VQKPLISNVEQLILGIPGQNRREIGHGQDIFPAEKLCHLQDRKVNLRHAA 111
278 WHKWIIAPKGFMANYPCHGECPFSLTISLNSSNYAFMQALMHAVDPEIPQ. 326
      | . | : || | . || | || . | | | : : | | | |
112 WGEIVAPKTLFSYSYCQGTCP.ALNSELRHSSF...ECYKRAV.PTCPWL 156
327 .AVCIPTKLSPISMLYQDNNDNVILRHIEDMVVDECGCG 364
      | || . | : : || . . . : . | : . || |
157 FQTCRPTMVRLFSLMVQDDEHKMSVHYVNTSLVEKCGCS 195

```

Percent Similarity: 32.941 Percent Identity: 26.471

FIG. 5A

```

tgagaaacac aatctgtatt atcaacttctt gcacctccat tctgtaaaca ggagttggta 60
ttgaagttgt tctgggagtg agagtttctc tcacttgaat ttaatttctc ttgaatgcgt 120
gatcagctac aagctgtggg ggggttagaat agggcctaca gctgggcacg tggatattta 180
aagacagcga aggggaagcc ccgcttctga gagcagggtat gttggagggg ggctgtggga 240
gaagtggcag ctcttggtc attcctgggc tcttggtctt gggctcttgg tgcattgtgt 300
tgagctcagt agagacgttt gactgtccca acccgatgct gccttcccac ataaatgaga 360
tttttttctg ccaggcaac atg gtt tta ccc tca tat tca aaa gtaagtagct 413
Met Val Leu Pro Ser Tyr Ser Lys 8
ggagcgctgg tctttgccag ggaaggagtg atccagaagc tgcctggcag cattttgtgg 473
ggctggtcag ggaatggggg gtaaatgaca acagatatta agggctcttg tgagtagagc 533
aaggagttgg gtacagaata ttcttcagct ggtctagcag aaatggaatc tgcttcctgg 593
tttcagctct gcaggcttgg tatgtaggat gtctttaagc tttatggctg atgccctaaa 653
gttctgtgtg taaggatgct ctaaagtgtg aagtacacag ctgctgggct gggcaactat 713
agtgttttgg gagataaaca gggcaagtgg cttgtcttag gtcattgtga ctggaatgat 773
tttcagtact agggcaatca ttctgactta attccagggg tagggtagtg ggagttgagg 833
aacctcagtc catccctggc tgctgtggac taagcactga ctttgacaag ctgagactgc 893
taagtctttg tcctgtcctg cccggctggg tagtggggag taagaagctg aaaggaggt 953
gggactttcc acgatagtgg cctcctggag cttccactct tctttcccta caggctcata 1013
gttcctacac agctactggc ttctctgttt tgaggcagtt tccttcttgg gggtttcctt 1073
gataaaagta tgggcttggg tgcccattgt ccccatgcc actgagcttg ttctagagtt 1133
cgaggaccat agaagggggc tccaaagatt ccttctggga tctttcccca ttatcttttc 1193
atcctaccag tcagagggag ggtcattatt ggatatctac tgtttactca cgtattggat 1253
ggaggtggtg cccaccctct tggcagagac aaagattcca gccactgatg tcgctgatgc 1313
cagcctgaat gaatgttcca gtaccgaaag gaaacaagac gtagtggtgc tgttcgtgac 1373
cttgtcccac acacagccac ctctgtttca cctgccttat gtccag aaa ccc tta 1428
Lys,Pro Leu 11
atc tct aat gtg gag cag ctg atc ctg ggg atc ccg ggc cag aat cgc 1476
Ile Ser Asn Val Glu Gln Leu Ile Leu Gly Ile Pro Gly Gln Asn Arg 27

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIG. 5B

cgg gag ata ggc cat ggc cag gat atc ttt cca gca gag aag ctc tgc	1524
Arg Glu Ile Gly His Gly Gln Asp Ile Phe Pro Ala Glu Lys Leu Cys	43
cat ctg cag gat cgc aag gtg aac ctt cac aga gct gcc tgg ggc gag	1572
His Leu Gln Asp Arg Lys Val Asn Leu His Arg Ala Ala Trp Gly Glu	59
tgt att gtt gca ccc aag act ctc agc ttc tct tac tgt cag ggg acc	1620
Cys Ile Val Ala Pro Lys Thr Leu Ser Phe Ser Tyr Cys Gln Gly Thr	75
tgc ccg gcc ctc aac agt gag ctc cgt cat tcc agc ttt gag tgc tat	1668
Cys Pro Ala Leu Asn Ser Glu Leu Arg His Ser Ser Phe Glu Cys Tyr	91
aag gtaagacatg gagcctcgtt ctttctcttc tggggtcata ttgggatagc	1721
Lys	92
actaagtgct caactctcta ggcttggtc cttttgagtc aaggaagcca ttgaagtgg	1781
taattatgta atctagcact gatgcagtgt gtagcatctt ccccgccctg tgaccttacc	1841
ccttatcttt attcataaga aacatcagct tcctaaagat tgttctgaaa cagccctgat	1901
ccagcagctt ctccccaggc cctccttctc ccttcccatg tatccctgac aagtctactg	1961
atgcccttag atatgaggct gtggctatga ggcactcacc attctgcat ttgtttctgc	2021
ag agg gca gta cct acc tgt ccc tgg ctc ttc cag acc tgc cgt ccc	2068
Arg Ala Val Pro Thr Cys Pro Trp Leu Phe Gln Thr Cys Arg Pro	107
acc atg gtc aga ctc ttc tcc ctg atg gtc cag gat gac gaa cac aag	2116
Thr Met Val Arg Leu Phe Ser Leu Met Val Gln Asp Asp Glu His Lys	123
atg agt gtg cac tat gtg aac act tcc ttg gtg gag aag tgt ggc tgc	2164
Met Ser Val His Tyr Val Asn Thr Ser Leu Val Glu Lys Cys Gly Cys	139
tct tga gatacccaaa agcctcctac tggcctcagg gccacctaag tctcaggact	2220
Ser *	140
ttagtagggg gtgggattac ttttcatagc aagtagagct ctttgaaggg aggtgggatt	2280
tggtttggtt ctcaaagcac agcaagaagg ttggcattat ggcaagtaacc cctcatagat	2340
gcttctcttt gatgtggcag gggcccccta gtgctgttct cagtcactcc tactactggg	2400
aagctgggcc cattgagatg tctgactatc gctgtcctag attgtgagtg ggctgggctt	2460
agtgccacct ctgggatcat ttaggtgggg aaagaggaac tggaattgga cgcagtgcag	2520
ctcttggggg aggggtaaaa ttgttaccag tgttaagctg gctttggact ctttctgagc	2580
cattcagctg ctatcatcct tctctgtacc attggcctgg ggctgggtcca gaactgacct	2640
cagcatgtac attcctcctc acctaacact cctggcctct ttagagggag tgaagactct	2700

FIG. 5B

FIG. 5C

gtggaagaaa gcattctgtc atgggctagt catgggtaaa gggccccaag gccttcacaa 2760
cctgggtgtca gatgggagcc tgagagtaga ggatgttgct tgactgacag agggggcctc 2820
tggcctcatg gaaagtttgt ctactatca tttaaggaac ttgatattag ctttttcact 2880
atctttaata aaactatagg accattgttg tgggtctctt atgttggata tctattactt 2940

gtggaagaaa gcattctgtc atgggctagt catgggtaaa gggccccaag gccttcacaa 2760
cctgggtgtca gatgggagcc tgagagtaga ggatgttgct tgactgacag agggggcctc 2820
tggcctcatg gaaagtttgt ctactatca tttaaggaac ttgatattag ctttttcact 2880
atctttaata aaactatagg accattgttg tgggtctctt atgttggata tctattactt 2940

[illegible]

tgagaaacac	aatctgtatt	atcactttct	gcacctccat	tctgtaaaca	ggagttggta	60
ttgaagttgt	tctgggagtg	agagttttct	tcacttgaat	ttaatTTTct	ttgaatgcgt	120
gatcagctac	aagctgtggg	gggttagaat	agggcctaca	gctgggcacg	tggatattta	180
aagacagcga	aggggaagcc	ccgcttctga	gagcaggtat	gttggagggt	ggctgtggga	240
gaagtggcag	ctcctggctc	attcctgggc	tcttggctct	gggtctttgg	tgcattgtgt	300
tgagctcagt	agagacgttt	gactgtccca	acccgatgct	gccttcccac	ataa atg	357
					Met	1
aga ttt ttt tct gcc agg caa cat ggt ttt acc ctc ata ttc aaa a						403
Arg Phe Phe Ser Ala Arg Gln His Gly Phe Thr Leu Ile Phe Lys						
gtaagtagc	tggagcgctg	gtctttgcc	gggaaggagt	gatccagaag	ctgcctggca	461
gcatttttgt	gggctggcca	gggaatgggg	tgtaaataac	aacagatatt	aagggctctt	522
gtgagtagag	caaggagttg	ggtacagaat	attcttcagc	tggtctagca	gaaatggaat	582
ctgcttctct	gtttcagctc	tgcaggcttg	gtatgttaga	tgtctttaag	ctttatggct	642
gatgccctaa	agttctgtgt	gtaaggatgc	tctaaagtgt	gaagtacaca	gctgctgggc	702
tgggcaacta	tagtgttttg	ggagataaac	agggcaagtg	gcttgtctta	ggcatgggtg	762
actggaatga	ttttcagtac	tagggcaatc	attctgactt	aattccaggg	gtaggggtgat	822
gggagttgag	gaacctcagt	ccatccctgg	ctgctgtgga	ctaagcactg	actttgacaa	882
gctgagactg	ctaagtcttt	gtcctgtcct	gcccggtgg	gtagtgggga	gtaagaagct	942
gaaagggagg	tgggactttc	cacgatagtg	gcctcctgga	gcttcactc	ttctttccct	1002
acaggctcat	agttcctaca	cagctactgg	cttctctgtt	ttgaggcagt	ttccttcttg	1062
ggggtttcct	tgataaagtt	atgggcttgg	gtgcccattg	tcccccattg	cactgagctt	1122
gttctagagt	tcgaggacca	tagaaggggc	ctccaaagat	tccttctggg	atctttcccc	1182
attatctttt	catcctacca	gtcagaggga	gggtcattat	tggatatcta	ctgtttactc	1242
acgtattgga	tggaggtggg	gcccaccctc	ttggcag ag	aca aag att cca gcc		1296
			Lys Thr Lys Ile Pro Ala			22
act gat gtc gct gat gcc agc ctg aat gaa tgt tcc agt acc gaa agg						1344
Thr Asp Val Ala Asp Ala Ser Leu Asn Glu Cys Ser Ser Thr Glu Arg						38
aaa caa gac gta gtg ttg ctg ttc gtg acc ttg tcc cac aca cag cca						1392
Lys Gln Asp Val Val Leu Leu Phe Val Thr Leu Ser His Thr Gln Pro						54

FIG. 6C

tctctgtacc attggcctgg ggctgggtcca gaactgacct cagcatgtac attcctcctc 2660
acctaacact cctggcctct ttagagggag tgaagactct gtggaagaaa gcattctgtc 2720
atgggctagt catgggtaaa gggccccaag gccttcacaa cctgggtgtca gatgggagcc 2780
tgagagtaga ggatgttgct tgactgacag agggggcctc tggcctcatg gaaagtttgt 2840
ctcactatca tttaaggaac ttgatattag ctttttcact atctttaata aaactatagg 2900
accattgttg tgggtctctt atgttggata tctattactt 2940

2940
2930
2920
2910
2900
2890
2880
2870
2860
2850
2840
2830
2820
2810
2800
2790
2780
2770
2760
2750
2740
2730
2720
2710
2700
2690
2680
2670
2660